Levels of agreement for RR intervals and short-term heart rate variability obtained from the Polar S810 and an alternative system

6 authors, including:

David Nunan
University of Oxford
120 PUBLICATIONS **2,287** CITATIONS

SEE PROFILE

Djordje Jakovljevic
Coventry University
122 PUBLICATIONS **1,726** CITATIONS

SEE PROFILE

Gay Donovan
Royal Brompton and Harefield NHS Foundation Trust
15 PUBLICATIONS **533** CITATIONS

SEE PROFILE

Lynette D Hodges
Massey University
25 PUBLICATIONS **773** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Special Issue "The Impact of Clinical Exercise Physiologists on Health and Wellbeing" (https://www.mdpi.com/journal/ijerph/special_issues/Clinical_Exercise_Physiologists) View project

Project   RELAX-CARDIO trial: The effect of Serelaxin on cardiac protection in chronic heart failure. View project

Levels of agreement for RR intervals and short-term heart rate variability obtained from the Polar S810 and an alternative system

David Nunan · Djordje G. Jakovljevic · Gay Donovan · Lynette D. Hodges · Gavin R. H. Sandercock · David A. Brodie

Abstract The aim of this study was to assess the agreement between HRV measures derived from a time series of RR intervals recorded by a standard 12-lead ECG (CP) and a commercially available RR interval recorder (S810). Thirty-three participants (19 males) (median age 36, range 20–63) underwent simultaneous, 5-min, supine RR-interval recordings. Each RR interval time series was analysed using the software supplied with the recording equipment. Two comparisons were then made. First, a comparison of RR interval data recording and editing only was made. Second, comparisons were made for measures of HRV derived from edited RR interval data. Agreement between RR intervals and standard HRV measures were assessed using intraclass correlation coefficient and limits of agreement. Agreement of HRV measures derived from RR intervals recorded and edited by individual systems was not acceptable. Agreement analyses for the number of RR intervals recorded and edited by each systems software showed excellent intraclass correlation coefficients (ICC lower 95% CI > 0.75) and acceptably narrow limits of agreement (LoA). These data indicate that the number of RR intervals recorded by S810 can agree well those recorded from a standard 12-lead ECG. This is true even after application of system specific data editing procedures. Commercial RR-interval recorders may offer a simple, inexpensive alternative to full 12-lead ECG in the recording and editing of RR intervals for subsequent HRV analysis in healthy populations.

Keywords RR interval numbers · Fast Fourier transformation · Autoregressive · Practical · Realistic · Interchangeable

Introduction

Heart rate variability (HRV) is seen as a vital non-invasive indicator of cardiovascular autonomic function in the analysis of physiological signals (Cerruti et al. 2006; Parati et al. 2006). Advances in telemetric technology have resulted in the introduction of ambulatory, wireless heart rate monitor (HRM) systems capable of RR interval recording (Polar S810). Combined with software developments (Polar Precision 4.03, Polar OY, Finland), such HRM technologies now provide a commercially available means of recording, editing and analysing RR interval data for short-term measures of HRV.

Independent assessment of new technologies is required to ensure validity of measures (Task Force 1996). Two studies have attempted to validate HRV measures obtained by the S810 (Kingsley et al. 2005; Gamelin et al. 2006). These studies, however, only assessed the validity of the S810 HRM in terms of RR interval recording. There are no studies that have assessed the editing and HRV analysis capabilities of the S810 software (Polar Precision) with any other commercial system.

D. Nunan (✉) · D. G. Jakovljevic · G. Donovan · D. A. Brodie
Research Centre for Society and Health,
Buckinghamshire New University, Gorelands Lane,
Chalfont St Giles, HP8 4AD, Buckinghamshire, UK
e-mail: dnunan01@bcuc.ac.uk

L. D. Hodges
Department of Sport and Exercise Science,
University of Bedfordshire, Bedfordshire, UK

G. R. H. Sandercock
Centre for Sports and Exercise Science,
University of Essex, Colchester, UK

The aim of this study was to assess the agreement between RR interval and HRV measures derived from a time series of RR intervals recorded by the S810 and a standard 12-lead ECG (CP). Agreement was assessed for RR-interval recordings and measures of HRV derived from each time series that were edited and analysed using each systems manufacturer-supplied software.

Methods

Study population

Thirty-three volunteers, 19 males with median age 34 (range 20–59) and 14 females with median age 48 (range 25–63), gave full written informed consent to participate in the study. The mean ± SD stature and mass for all participants was 1.73 ± 0.11 m and 74.6 ± 15.6 kg. All participates were in good health as defined by the absence of cardiovascular disease and were not taking any medication that may have influenced HRV during the study period. All procedures were approved by the local ethics committee and conformed to the declaration of Helsinki.

Instrumentation and data acquisition

RR intervals were recorded simultaneously via the CardioPerfect (CP) software (Cardio Control, Delft, The Netherlands) within the Medical Graphics CardiO$_2$ stress system (Medical Graphics Corporation, St Paul, MN, USA) and a Polar S810 HRM (Polar Electro OY, Kempele, Finland). The CP uses a 12-lead ECG configuration and the S810 consists of a chest strap transmitter plus wrist watch receiver. The CP system was chosen as the criterion measure of HRV as its precision of measurement and reliability has been established (Sandercock et al. 2004a, b). Both CP and S810 were set to record with the same sampling frequency of 1,000 Hz, providing a temporal resolution of 1 ms for each RR period (Cottin et al. 2004).

The sampling time for CP recordings was set at 300 s and the digitised signal was stored on the hard drive of a PC (Dell Precision 340, Dell Computers, TX, USA). The RR interval data from CP recordings were edited and subsequently analysed for variability using the automated editing and HRV features of the CP software. The S810 recorded continuously for the duration of the CP recordings. S810 RR interval recordings were transferred to a password protected PC via the Polar Precision Performance 4.03 software. The Polar Precision software was used for both editing and HRV analysis of S810 RR interval recordings.

Experimental design

Effort was made to ensure all procedures were carried out in a quiet laboratory. Each participant was asked to abstain from caffeine and alcohol containing foods and beverages on test days and to avoid heavy physical exertion and alcohol consumption during the 48 h preceding test days.

Participants attended the laboratory between 8:00 and 13:00 on three occasions. On reporting to the laboratory, participants were prepared for recordings. This involved the cleaning and preparation of the skin for the attachment of surface electrodes (Blue Sensor Medicotest, Olstykke, Denmark). Electrodes were placed in the standard configuration for a 12-lead ECG. The S810 chest strap was correctly placed in accordance with the manufacturers guidelines.

Participants next lay on a bed while the signals from the equipment were checked for interference and signal quality and to allow heart rate to stabilise. Each participant then underwent two sequential 300 s CP recordings and one ∼10 min S810 recording. At the end of the first CP recording, participants were instructed to remain in the same position. The first CP recording was saved to the hard drive of the PC. During this time the S810 was still recording, but these data were not entered into the HRV analysis. Following saving of the first CP recording (∼30 s), a second CP recording was then performed. The start and end of the each CP recording was synchronised with the S810 recording using the temporal "event" (lap) marker feature of the S810. This procedure provided S810 recordings totalling approximately 10 min and 30 s. The number of RR intervals (RR count) returned by the S810 and CP systems, following editing of the time series, was obtained from each trial for subsequent agreement analysis.

Heart rate variability analysis

To ensure as stationary a signal as possible, the analysis of variation in RR intervals was performed on data from the second CP and corresponding S810 recordings. Prior to HRV analysis, raw RR intervals from both the CP and S810 recordings were edited and compared to discriminate error caused by S810 acquisition or by artefacts. Editing of CP and S810 recording artefacts was performed using default automated protocols within CP and Polar software programmes. For the CP system, abnormal intervals were defined as any interval differing by more than 20% from the previous interval. The Polar software corrects for artefacts using an error filter and beat protection zone function. A moderate filter power set at a minimum beat protection zone of six beats min^{-1} was used. Following abnormal interval removal, both the CP and Polar software linearly interpolated removed intervals using system specific algorithms. Standard time and frequency domain measures of

HRV were then derived from the normal-to-normal (NN) CP and S810 interval data within the software programme for each system.

In accordance with current recommendations (Task Force 1996) only the standard deviation of NN intervals (SDNN) and the root mean square of successive differences (RMSSD) were derived from time domain analysis. Mean RR interval was also obtained as a further index of cardiac autonomic control (Pinna et al. 2007).

By default the CP and S810 software use fast Fourier transformation (FFT) and autoregressive (AR) methods, respectively, to derive frequency domain HRV measures from the RR interval time series. In accordance with Task Force recommendations, the power spectrum for frequency domain HRV analysis was divided into the following bands for all three systems: very low frequency (VLF 0.0033–0.04 Hz), low frequency (LF 0.04–0.15 Hz) and high frequency (HF 0.15–0.40 Hz). Only LF, HF in absolute and normalised (LFnu, HFnu) units and LF:HF ratio were assessed for agreement.

Statistical analysis

All statistical analysis was carried out using SPSS version 13.0 (SPSS inc., Chicago, IL, USA). Normality of distribution was assessed using a Kolmogorov–Smirnov test. Where normality assumptions were not met, data were subsequently log-transformed.

To assess differences in RR count between the S810 and CP systems, repeat measures ANOVA were performed. The number of RR intervals recorded by the two systems was assessed for agreement by intraclass correlation coefficient (ICC) with 95% CI. An ICC of >0.80 is commonly considered as indicating good to excellent agreement and ICCs between 0.60 and 0.80 are taken to represent substantial agreement (Pinna et al. 2007). However, agreement sufficient for the interchangeable use of two methods is suggested only when a lower 95% CI value of >0.75 is observed (Lee et al. 1989; Sandercock et al. 2004b).

Bland and Altman (1986) query the use of a single number to summarise agreement and suggest the calculation of 95% CI based on the mean of differences between two methods in addition to the ICC. Plots of average values from both systems versus differences between systems and subsequent limits of agreement (LoA) are recommended. Therefore, 95% LoAs were calculated for RR intervals recorded and edited by the S810 and CP systems. The recently described approach by Bland and Altman (2007) was adopted for LoA analysis due to the multiple observations per individual in the present study.

Repeat measures ANOVAs were carried out on all HRV measures to assess systematic bias between the two systems (Bland and Altman 1986, 1990; Hopkins 2000). As with the

number of RR intervals, agreement between measures of HRV derived by the two systems was assessed using ICC and LoA.

For LoA analysis of log-transformed data, dimensionless ratios were calculated by taking the antilog of the mean of differences and 95% LoA (Bland and Altman 1986). Outcomes of transformed data are presented and described separately to those of non-transformed data as recommended (Mortensen et al. 2002).

Results

A technical failure with the CP excluded the first trial data of two participants, the second trial data of one participant and the third trial data of eight participants. As a result, data from all three trials were only available for 23 participants. Data from two trials were obtained for nine participants and for one participant data were only available from one trial. Statistical outliers were found for the measures LF and HF. These values, however, were not removed as they were apparent in both systems and were also within the range of reported values for these measures.

There were no differences in the number of RR intervals recorded by the S810 in comparison to the CP for any trial (Table 1). There was agreement sufficient to allow interchangeable use of the two systems in the recording and editing of RR intervals, with an average ICC of 0.97 and a lower bound CI >0.75 across all three trials. 95% LoA revealed the number of RR intervals returned by the S810 following editing is between 22 less than and 25 more than CP, with the S810 returning on average one more RR interval than that of the CP (Table 3; Fig. 2a).

All measures of HRV except mean RR interval time, LFnu, HFnu and LF:HF showed a marked right-skewed distribution (Kolmogorov–Smirnov test $P < 0.05$). Subsequent log-transformation of skewed measures provided normality and homoscedasticity (Fig. 2). Descriptive statistics of non-transformed HRV measures and their agreement as assessed by ICC are reported in Table 1.

There were no significant differences between values for non-transformed measures of HRV. All four measures showed excellent levels of agreement when assessed by ICC (Table 1) but only mean RR interval displayed sufficient agreement across all three trials to allow each system to be used interchangeably (lower 95% CI >0.75). However, analysis by LoA revealed that mean RR, LFnu, HFnu and LF:HF displayed unacceptable agreement between systems (Table 3).

There was no systematic bias between the two systems for log-transformed measures of HRV (Table 2). Analysis by ICC showed excellent and interchangeable agreement for SDNN and RMSSD across all three trials and for LF

Table 1 Descriptive results and agreement outcomes from analysis by intraclass correlation coefficient for homoscedastic HRV measures obtained by the S810 and CP

		S810	CP	*P**	ICC (95% CI)	Interchangeable agreement
RR count	Trial 1	316 ± 54	314 ± 53	0.99	0.98 (0.95–0.99)	Yes
	Trial 2	320 ± 53	319 ± 55		0.96 (0.92–0.98)	Yes
	Trial 3	299 ± 40	298 ± 39		0.98 (0.97–0.99)	Yes
Mean RR (ms)	Trial 1	980.6 ± 178.6	979.4 ± 176.9	0.99	0.99 (0.98–0.99)	Yes
	Trial 2	970.3 ± 167.9	964.7 ± 172.6		0.98 (0.95–0.99)	Yes
	Trial 3	1021.8 ± 140.9	1026.3 ± 136.8		0.98 (0.96–0.99)	Yes
LF (nu)	Trial 1	62.5 ± 14.5	59.0 ± 17.8	0.59	0.75 (0.54–0.87)	No
	Trial 2	62.9 ± 14.5	58.5 ± 18.6		0.73 (0.51–0.86)	No
	Trial 3	62.2 ± 16.0	61.3 ± 16.3		0.70 (0.44–0.87)	No
HF (nu)	Trial 1	37.5 ± 14.5	41.3 ± 17.7	0.43	0.72 (0.51–0.86)	No
	Trial 2	37.1 ± 14.5	42.1 ± 18.7		0.71 (0.48–0.85)	No
	Trial 3	37.8 ± 16.0	38.2 ± 15.9		0.74 (0.49–0.87)	No
LF:HF	Trial 1	2.2 ± 1.9	2.1 ± 2.1	0.72	0.90 (0.81–0.95)	Yes
	Trial 2	2.2 ± 1.5	2.1 ± 1.9		0.87 (0.76–0.94)	Yes
	Trial 3	2.1 ± 1.3	2.3 ± 1.9		0.54 (0.19–0.76)	No

Data represent the mean ± SD

RR count number of RR interval data points recorded, *mean RR* mean time between normal r-waves, *LF* low frequency spectral power, *HF* high frequency spectral power, *LF:HF* the ratio of low to high frequency spectral power; *nu* normalised units, *ICC* intraclass correlation coefficient, *CI* confidence interval

* *P* value outcomes from repeated measures ANOVA for group effect

Table 2 Descriptive results and agreement outcomes from analysis by intraclass correlation coefficient for heteroscedastic HRV measures obtained by the S810 and CP

		S810	CP	*P**	ICC (95% CI)	Interchangeable agreement
Ln SDNN (ln ms)	Trial 1	4.06 ± 0.49	4.01 ± 0.53	0.66	0.87 (0.75–0.94)	Yes
	Trial 2	4.10 ± 0.47	4.11 ± 0.45		0.94 (0.88–0.97)	Yes
	Trial 3	4.02 ± 0.41	4.09 ± 0.41		0.97 (0.92–0.99)	Yes
Ln RMSSD (ln ms)	Trial 1	3.70 ± 0.61	3.68 ± 0.67	0.81	0.88 (0.77–0.94)	Yes
	Trial 2	3.75 ± 0.62	3.82 ± 0.70		0.88 (0.77–0.94)	Yes
	Trial 3	3.77 ± 0.58	3.88 ± 0.63		0.94 (0.87–0.97)	Yes
Ln LF (ln ms^2)	Trial 1	6.93 ± 0.92	6.58 ± 1.24	0.23	0.64 (0.38–0.81)	No
	Trial 2	6.92 ± 0.91	6.82 ± 1.09		0.84 (0.70–0.92)	No
	Trial 3	6.92 ± 0.88	6.87 ± 0.92		0.89 (0.79–0.95)	Yes
Ln HF (ln ms^2)	Trial 1	6.37 ± 1.17	6.20 ± 1.31	0.81	0.81 (0.65–0.91)	No
	Trial 2	6.34 ± 1.19	6.41 ± 1.42		0.93 (0.87–0.97)	Yes
	Trial 3	6.38 ± 1.16	6.50 ± 1.10		0.81 (0.62–0.91)	No

Data represent the mean ± SD

SDNN the standard deviation of normal-to-normal intervals, *RMSSD* root mean square of successive differences, *LF* low frequency spectral power, *HF* high frequency spectral power, *ln* natural logarithm, *ICC* intraclass correlation coefficient, *CI* confidence interval

* *P* value outcomes from repeated measures ANOVA for group effect

trial three and HF trial two. Good to excellent but not interchangeable agreement was found for LF and HF in their other two trials.

LoA analysis for heteroscedastic HRV measures are reported in Table 4. Since statistical analysis of these vari-

ables was carried out following log-transformation, the bias and 95% LoA are expressed in log units. To relate these data to the original scale of measurement, antilogarithmic transformation provides dimensionless ratios where a ratio of one is equal to zero. Measurement values from the S810

Table 3 Outcome of limits of agreement analyses between homoscedastic heart rate variability (HRV) measures obtained by the S810 and CP

	Comparison	Bias	Limits of agreement	Acceptable limits
RR count	S810 versus CP	1.4	±23.2	Yes
Mean RR	S810 versus CP	2.5 ms	±61.8 ms	No
LF (nu)	S810 versus CP	3.1 ms	±23.2. ms	No
HF (nu)	S810 versus CP	−3.3 ms	±23.4 ms	No
LF:HF	S810 versus CP	0.06	2.23	No

The legend for HRV measures is given in Table 1. Comparisons are made between values from the former system to values from the second system (e.g. S810 vs. CP = values obtained by the CP are subtracted from values obtained by S810). Bias = the mean of differences in participant values between the two systems for each test (three tests $n = 23$, two tests $n = 9$, one test $n = 1$, total tests $n = 88$)

system can be as small/large as about 0.3/4.4 times those of the CP; with S810 measures averaging 1–1.2 times those of the CP.

Discussion

The research and clinical studies assessing cardiac autonomic activity has led to the development of more accessible and user friendly technologies and systems to determine heart rate variability (HRV). One such system is the Polar S810 HRM which provides an alternative means to determine HRV away from the laboratory setting. To ensure comparability between laboratory and field measures of HRV it is important to quantify the agreement between methods and systems being used. In this study, we carried out an in-depth assessment of the agreement between measures of HRV obtained by the S810 and by a 12-lead ECG

laboratory based system (CP). Despite substantial to excellent agreement of HRV measures when assessed by ICC, analysis by LoA revealed measures of HRV from the S810 agree poorly with those from the CP system. When obtained using default settings, measures could vary considerably and unacceptably. However, the number of RR intervals obtained from edited S810 recordings showed excellent and interchangeable agreement with that obtained from CP system recordings.

There were no significant differences in values for non-transformed or transformed measures of HRV, indicating an absence of systematic bias between the two systems. Analysis by ICC showed that, at worst, agreement between the S810 and CP systems was substantial (LF:HF, trial 3) and at best excellent. All time domain measures display interchangeable levels of agreement across all trials and three frequency domain measures (LF, HF and LF:HF) display interchangeable agreement in at least one trial. However, the sole use of ICC to demonstrate agreement has been questioned, primarily due to fact that values can be exaggerated when data reveal a wide spread of scores, often disguising the true magnitude of variation (Bland and Altman 1990). Large SD values for the majority of HRV measures (Table 1) indicated large inter-individual variations, a finding not uncommon in HRV analysis (Sinnreich et al. 1998; Pikkujamsa et al. 2001). Such wide spreads of scores can inflate the values for ICC coefficients and may disguise the true magnitude of variation.

Analysis using the Bland and Altman (2007) limits of agreement (LoA) method revealed that agreement between the systems was poor, particularly for absolute measures in the frequency domain (see Tables 3, 4). On average and for the majority of HRV measurements, the S810 showed a bias for higher values compared to those of the CP. Exceptions to this were found for measures of normalised HF and

Table 4 Outcome of limits of agreement analyses between heteroscedastic heart rate variability (HRV) measures obtained by the S810 and CP

	Comparison	Bias	Upper, lower 95% LoA	Antilog of bias	Antilog of upper, lower 95% LoA	Interpretation of antilog values	
						Bias	Upper, lower 95% LoA
Ln SDNN	S810 versus CP	0.00 ln ms	−0.38, 0.36	0.00	0.68, 1.45	On average S810 yields the same value as CP	S810 may yield between 0.68 and 1.45 times that of CP
Ln RMSSD	S810 versus CP	−0.05 ln ms	−0.61, 0.51	0.95	0.54, 1.66	On average S810 yields 0.95 times CP	S810 may yield between 0.54 and 1.66 times that of CP
Ln LF	S810 versus CP	0.18 ln ms^2	−1.13, 1.49	1.19	0.32, 4.44	On average S810 yields 1.19 times CP	S810 may yield between 0.32 and 4.44 times that of CP
Ln HF	S810 versus CP	0.00 ln ms^2	−1.29, 1.29	0.00	0.28, 3.63	On average S810 yields the same value as CP	S810 may yield between 0.28 and 3.63 times that of CP

The legend for HRV measures is given in Table 2. Comparisons are made between values from the former system to values from the second system (e.g. S810 vs. CP = values obtained by the CP are subtracted from values obtained by S810). Bias = the mean of differences in participant values between the two systems for each test (three tests $n = 23$, two tests $n = 9$, one test $n = 1$, total tests $n = 88$). The antilog values are a dimensionless ration, where 1 = zero. A value less than 1 represents a negative bias from one system compared to the other. A value greater than 1 represents a positive bias from one system to the other
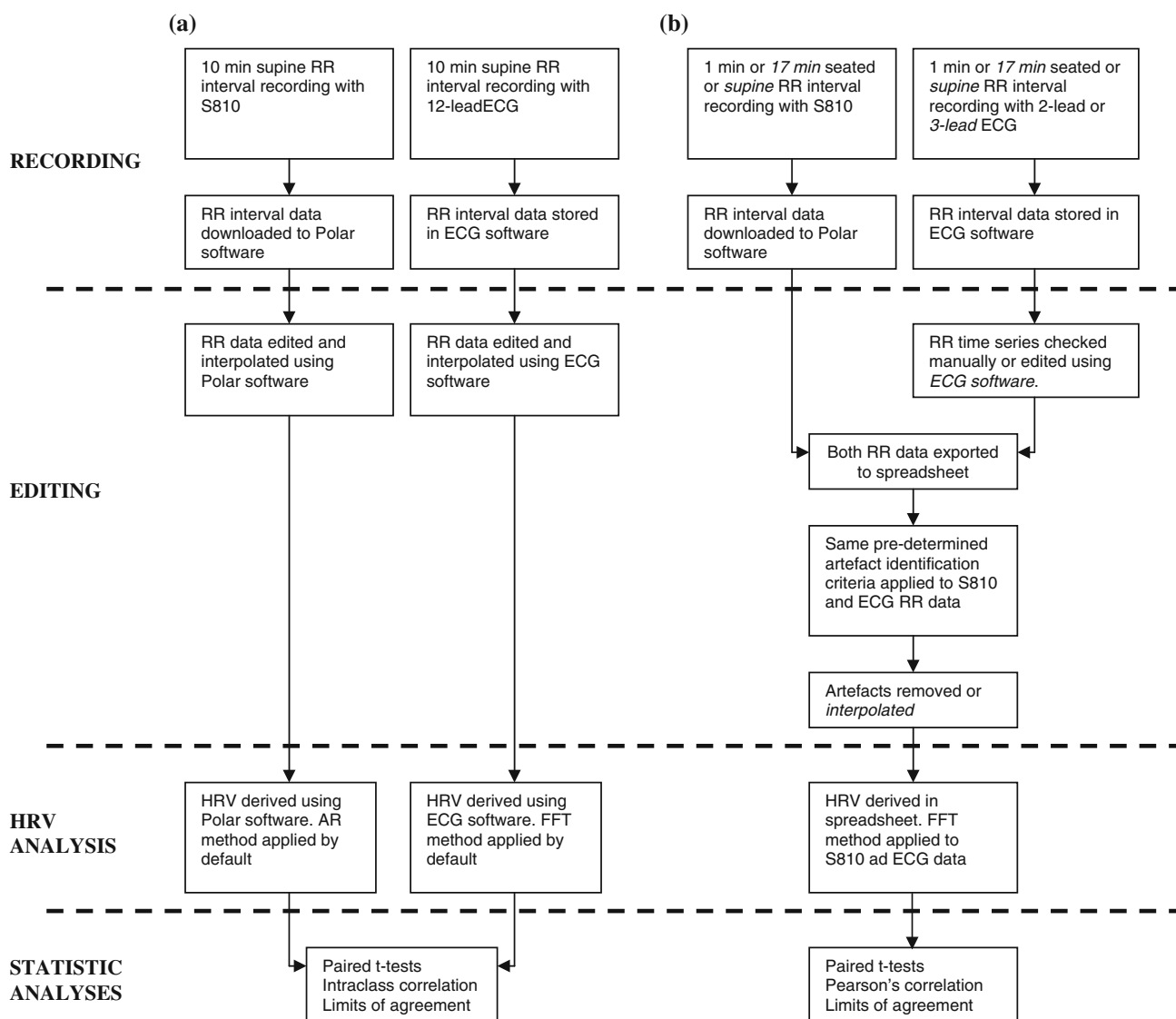
(a)

10 min supine RR interval recording with S810	10 min supine RR interval recording with 12-leadECG

RECORDING

RR interval data downloaded to Polar software	RR interval data stored in ECG software

- -

RR data edited and interpolated using Polar software	RR data edited and interpolated using ECG software

EDITING

- -

HRV derived using Polar software. AR method applied by default	HRV derived using ECG software. FFT method applied by default

HRV ANALYSIS

- -

STATISTIC ANALYSES

Paired t-tests
Intraclass correlation
Limits of agreement

(b)

1 min or *17 min* seated or *supine* RR interval recording with S810	1 min or *17 min* seated or *supine* RR interval recording with 2-lead or *3-lead* ECG

RR interval data downloaded to Polar software	RR interval data stored in ECG software

- -

RR time series checked manually or edited using *ECG software.*

Both RR data exported to spreadsheet

Same pre-determined artefact identification criteria applied to S810 and ECG RR data

Artefacts removed or *interpolated*

- -

HRV derived in spreadsheet. FFT method applied to S810 ad ECG data

- -

Paired t-tests
Pearson's correlation
Limits of agreement

Fig. 1 Procedures adopted by the present study (**a**) and those adopted by Kingsley et al. (2005) and Gamelin et al. (2006) (**b**) for the recording, editing and analysis of S810 and ECG RR interval data. *Italicised* words are representative of procedures carried out by Gamelin et al. (2006)

RMSSD. As an example, when obtained by the S810 the value for a simple measure such as mean RR interval could vary by as much as plus or minus 62 ms compared to that of CP (Fig. 2b). A more extreme case is presented for frequency domain measures such as LF, where values from the S810 could be anywhere from one-third to four times those obtained from the CP system. These findings suggest that the S810 and CP systems should not be used interchangeably in the determination of commonly used HRV measures.

The bias presented by S810 can be explained by small discrepancies in the length of RR intervals and differences in frequency domain analysis methods between the CP and S810 systems. The use of differing algorithms for detecting

R-wave peaks may have caused the small discrepancies in RR interval length observed between the S810 and CP systems and these were likely to have been further magnified as a result of the calculations performed to assess variability.

The CP system utilises the FFT method to determine frequency domain measures of HRV. By contrast, the S810 relies on autoregressive (AR) methods. Although these methods can create the same measures, values may differ (Fagard 2001; Chemla et al. 2005; Pichon et al. 2006). Differences in estimates of spectral power outcomes from FFT compared to those estimated by AR methods have been related to the crossover of power between defined bandwidths (Badilini et al. 1998) and the possible inclusion

Fig. 2 Bland Altman plots for RR count **a**, mean RR **b**, log-transformed SDNN **c**, RMSSD **d**, LF **e**, HF **f** absolute measures of HRV
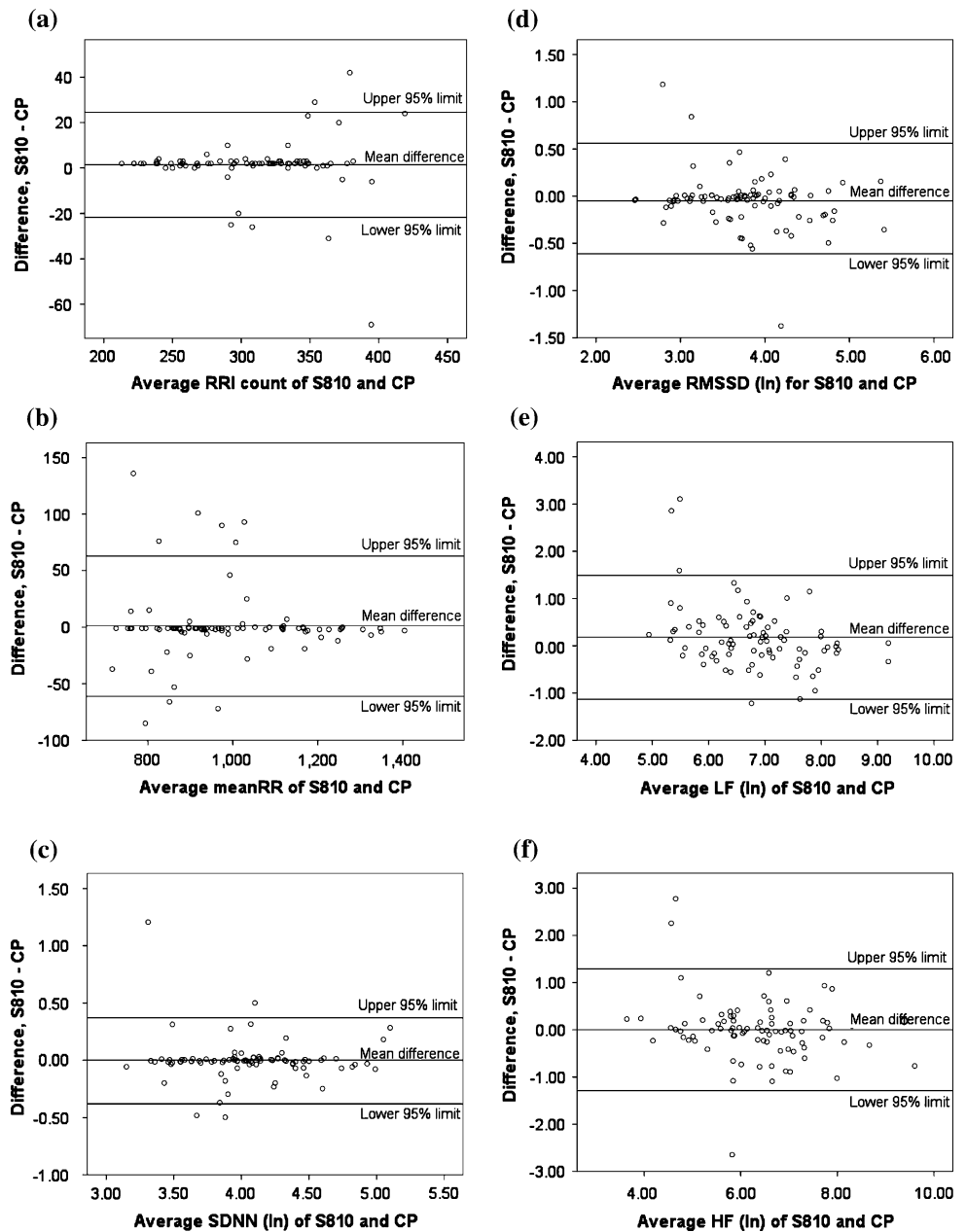
of noise with the FTT method (Fagard et al. 1998). Finally, the use of the elastic electrode belt for the S810 system may induce minor artefacts (Gamelin et al. 2006).

Comparisons with the present study findings are difficult due to the lack of studies assessing agreement of short-term HRV measurements, differences in data collection protocols and the use of inappropriate or insufficient statistical tests. That said, the findings of unacceptable agreement presented here are in agreement with those of Carrasco et al. (1998). In assessing the agreement of commonly used HRV measures obtained from short ECG and blood pressure wave recordings, the Carrasco et al. study found a systematic underestimation in resting and exercising blood pressure wave measures. Although some time domain measures

displayed good LoA under certain conditions, these were the exceptions to the finding of generally poor LoA, particularly for frequency domain measures.

In a more recent study assessing short-term measures of HRV from three separate instruments capable of recording and analysing ECG data, Sandercock et al. (2004b) reported similar and unacceptable levels of agreement. A similar finding for excellent agreement was found when assessed by ICC analysis. Likewise, LoA analysis revealed that variation between systems could be large for any measure of HRV. By assessing for expected effect sizes of a few select measures, Sandercock et al. reported that interchangeable use of the systems could mask differences between groups and/or prevent the detection of changes

following exercise interventions. Values for expected differences in HF power of 704 ms^2 between trained and untrained subjects and in LFnu of 2.3 following exercise training were obtained from Aubert et al. (2001) and Leicht et al. (2003). Using these data, possible variations in raw HF of $\pm 2,545$ ms^2 and LFnu of ± 23.2 between the S810 and CP would make identifying differences between groups and/or changes due to intervention extremely difficult if the two systems were used interchangeably.

The present study findings show that HRV measures derived from S810 RR interval data and Polar specific software appear not to agree with those derived from a 12-lead ECG. However, the number of RR intervals recorded by the two systems showed excellent and interchangeable agreement. Importantly, this was found after editing of raw RR interval data by each systems specific software programmes. These findings are encouraging as they demonstrate the ability of a two lead, telemetric ECG recording system to create comparable RR interval data compared to a 12-lead ECG. Previous studies attempting to assess RR interval data obtained by the S810 have shown different outcomes to those of the present study, not least due to their employed RR interval editing procedures (Kingsley et al. 2005; Gamelin et al. 2006). This point is illustrated better in Fig. 1. It shows the differences between RR interval data editing and HRV analysis procedures performed in the present study and in those of Kingsley et al. (2005) and Gamelin et al. (2006). The important point in Fig. 1 is that neither of the previous studies assessed the RR recording or HRV analysis abilities of the S810 with its own designated software. Exporting S810 and ECG RR data to the same spreadsheet and applying the same editing, interpolation, re-sampling, de-trending and HRV analysis procedures to both data sets, will produce situations that are unrealistic for the more practical use of the S810. Moreover, the employment of such a methodology is likely to result in greater similarities of RR interval data and explains the extremely high levels of agreement for measures of HRV reported by both studies.

Only Gamelin et al. (2006) present data for the number of RR intervals recorded by the S810 and ECG system but neither study assessed differences or agreement between the number of RR intervals recorded by each system. In addition, by performing analyses on all RR intervals recorded for each individual (i.e. 300 beats from one individual repeated over n number of individuals), we conclude that both studies wrongly assessed both inter-and intra-individual agreement from the same LoA analyses (Bland and Altman 1986).

This paper demonstrates an example of a more practical and realistic scenario in which RR intervals and HRV measures are likely to be derived using the S810 and shows that S810 recordings can provide RR intervals comparable in

number to those of a 12-lead ECG. Users should be aware, however, that HRV measures derived from factory default settings from different systems may yield widely varying outcomes and using these systems interchangeably to measure HRV is not recommended.

Conclusions

This study assessed the agreement of HRV measurements derived from the S810 and accompanying Polar software. When derived in this manner, and despite high ICCs, measures of HRV obtained by the S810 display unacceptable agreement with those simultaneously obtained by the 12-lead ECG (CP) system. Non significant differences, interchangeable agreement by ICC and narrow LoA for the number of RR intervals recorded, suggest that S810 and CP are comparable in their ability to record and process RR interval data. The two systems may, therefore, be used interchangeably in the recording and, where appropriate, the interpolation of RR intervals.

References

Aubert AE, Beckers F, Ramaekers D (2001) Short-term heart rate variability in young athletes. J Cardiol 37(suppl 1):85–89

Badilini F, Maison-Blanche P, Coumel P (1998) Heart rate variability in passive tilt test: comparative evaluation of autoregressive and FFT spectral analyses. Pacing Clin Electrophysiol 21:1122–1132

Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 8476:307–310

Bland JM, Altman DG (1990) A note on the use of the intraclass correlation coefficient in the evaluation of agreement between two methods of measurement. Comput Biol Med 20(5):337–340

Bland JM, Altman DG (2007) Agreement between methods of measurement with multiple observations per individual. J Biopharmaceut Stat 17:571–582

Carrasco S, Gonzalez R, Jimenez J, Roman R, Medina V, Azpiroz J (1998) Comparison of the heart rate variability parameters obtained from the electrocardiogram and the blood pressure wave. J Med Eng Technol 22(5):195–205

Cerruti S, Golderberer AL, Yamamoto Y (2006) Recent advances in heart rate variability signal processing and interpretation. IEEE Trans Biomed Eng 55(1):1–3

Chemla D, Young J, Badilini F, Maison-Blanche P, Affres H, Lecarpentier Y, Chanson P (2005) Comparison of fast Fourier transform and autoregressive analysis in the study of heart rate variability in diabetic patients. Int J Cardiol 104:307–313

Cottin F, Médigue C, Leprêtre P-M, Papelier Y, Koralsztein J-P, Billat V (2004) Heart rate variability during exercise performed below and above ventilatory threshold. Med Sci Sport Exerc 36(4):594–600

Fagard RH (2001) A population-based study on the determinants of heart rate and heart rate variability in the frequency domain. Verh Koninklijke Acad Geneeskd Belg 63(1):57–89 (discussion 90–51)

Fagard RH, Pardaens K, Staessen JA et al (1998) Power spectral analysis of heart rate variability by autoregressive modelling and fast Fourier transform: a comparative study. Acta Cardiol 53:211–218

Gamelin FX, Berthoin S, Bosquet L (2006) Validity of the polar S810 heart rate monitor to measure RR intervals at rest. Med Sci Sport Exerc 38:887–893

Hopkins WG (2000) Measures of reliability in sports medicine and science. Sports Med 30(1):1–15

Kingsley M, Lewis MJ, Marson RE (2005) Comparison of polar 810s and an ambulatory ECG system for RR interval measurement during progressive exercise. Int J Sports Med 26:39–44

Lee J, Koh D, Ong CN (1989) Statistical evaluation of agreement between two methods for measuring a quantitative measure. Comput Biol Med 19:61–70

Leicht AS, Allen GD, Hoey AJ (2003) Influence of age and moderate-intensity exercise training on heart rate variability in young and mature adults. Can J Appl Physiol 28(3):446–461

Mortensen S, Lose G, Thyseen H (2002) Repeatability of cystometry and pressure-flow parameters in female patients. Int Urogynecol J 13:72–75

Parati G, Mancia G, Di Rienzo M, Castiglioni P Jr, Talyor A, Studinger P (2006) Point:counterpoint: cardiovascular variability is/is not an index of automatic control of circulation. J Appl Physiol 101:676–682

Pichon A, Roulard M, Antoine-Jonville S, de Bisschop C, Denjean A (2006) Spectral analysis of heart rate variability: interchangeability between autoregressive analysis and fast Fourier transform. J Electrocardiol 39:31–37

Pikkujamsa SM, Makikallio TH, Airaksinen KEJ, Huikuri HV (2001) Determinants and interindividual variation of R–R interval dynamics in healthy middle-aged subjects. Am J Physiol Heart C 280(3):H1400–H1406

Pinna GD, Maestri R, Torunski A, Danilowicz-Szymanowiczm L, Szwochm M, La Rovere MT, Raczak G (2007) Heart rate variability measures: a fresh look at reliability. Clin Sci (London) 113(3):131–140

Sandercock GRH, Bromley PD, Brodie DA (2004a) Reliability of three commercially available heart rate variability instruments using short-term (5-min) recordings. Clin Physiol Funct Imaging 24:359–367

Sandercock GRH, Shelton C, Bromley P, Brodie DA (2004b) Agreement between three commercially available instruments for measuring short-term heart rate variability. Physiol Meas 25:1115–1124

Sinnreich R, Kark JD, Friedlander Y, Sapoznikov D, Luria MH (1998) Five minute recordings of heart rate variability for population studies: repeatability and age–sex characteristics. Heart 80(2):156–162

Task Force (1996) Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation 93:1043–1065